Huub van Dam

Application Architect at Brookhaven National Laboratory

huubvandam@aol.com

Summary

Computational chemist experienced in writing and supporting large parallel quantum chemistry packages. Past experience has focussed on method development for describing electron correlation effects by either wavefunction methods such as multi-reference perturbation theory or electron density based methods such as density functional theory, as well as properties based on these energy expressions. Most recent work has focussed on taking these capabilities to extreme scale parallel computing platforms and addressing in particular the fault tolerance aspects of this.

Interested in developing methods that are quantitatively predictive as well as fast for a wide range of applications.

Specialties: Density functional theory code development,
Wavefunction methods development including multi-reference perturbation theory, configuration interaction, and multi-configurational self consistent field methods,
Fault tolerance in HPC applications,
Global array and MPI based parallel programming

Experience

Application Architect at Brookhaven National Laboratory
September 2015 - Present (1 year 4 months)

Scientist at Environmental Molecular Science Laboratory at Pacific Northwest National Laboratory
October 2009 - August 2015 (5 years 11 months)

Development of quantum chemistry methods in NWChem. This relates to the density functional theory as well as the extreme scale parallel computing capabilities. It also extends to associated activities such as benchmark development, user support, and novel application projects.

Higher Scientific Officer at Science and Technology Facilities Council
August 1998 - September 2009 (11 years 2 months)

My work involved largely the ongoing development of GAMESS-UK, an electronic structure code. I worked on aspects of the parallisation and performance analysis, including adding Vampirtrace instrumentation to the Global Array library, as well as new capabilities such as RPA gradients, DFT hessians, as well as generally hardening up the DFT capabilities, including developing a library of DFT functionals for general use that supports second partial derivatives. User support, software licensing and training were also part of my duties.

PhD Student at Utrecht University

1993 - 1997 (4 years)

My research for my PhD involved mainly theory and method development for post Hartree-Fock calculations of molecular electronic structure. In particular I focused on the question whether multi-reference Moller-Plesset perturbation theory can be formulated in a size consistent way. I showed that dependent on the choice of zeroth-order Hamiltonian this can indeed be done.

Languages

Dutch
German
English

Skills & Expertise

Fault Tolerance
Parallel Programming
Quantum Chemistry
High Performance Computing
MPI
Parallel Computing
Theory
Scientific Computing
Computational Chemistry
Density Functional Theory
Chemistry
OpenMP
Physical Chemistry
Fortran
Molecular Modeling
Computational Physics
Physics
Supercomputing
Nanomaterials
Spectroscopy
Numerical Analysis
Materials Science
Computational Biology
Nanotechnology
Mathematical Modeling
Surface Chemistry
LaTeX
Science
Catalysis

Experimentation
Numerical Simulation
Organometallic Chemistry
Structural Biology
Bioinformatics
Cheminformatics

Publications

Structure, optical properties and defects in nitride (III-V) nanoscale cage clusters

Physical Chemistry and Chemical Physics 2008

Authors: Huub van Dam, Stephen A. Shevlin, Z. Xiao Guo, Paul Sherwood, C. Richard A. Catlow, Alexey A. Sokol, Scott M. Woodley

Density Functional Theory calculations are reported on cage structured BN, AlN, GaN and InN sub- and low nanosize stoichiometric clusters, including two octahedral families of T_d and T_h symmetry. The structures and energetics are determined, and we observe that BN clusters in particular show high stability with respect to the bulk phase. The cluster formation energy is demonstrated to include a constant term that we attribute to the curvature energy and the formation of six tetragonal defects. The (BN)_60 onion double-bubble structure was found to be particularly unstable. In contrast, similar or greater stability was found for double and single shell cages for the other nitrides. The optical absorption spectra have been first characterised by the one-electron Kohn-Sham orbital energies for all compounds, after which we concentrated on BN where we employed a recently developed Time Dependent Density Functional Theory approach. The one-electron band gaps do not show a strong and consistent size dependency, in disagreement with the predictions of quantum confinement theory. The density of excited bound states and absorption spectrum have been calculated for four smallest BN clusters within the first ionisation potential cut-off energy. The relative stability of different BN clusters has been further explored by studying principal point defects and their complexes including topological B-N bond rotational defects, vacancies, antisites and interstititials. The latter have the lowest energy of formation.

Is MPI-2 suitable for Quantum Chemistry? Performance of passive target one-sided communications

HPCx Technical Reports 2008

Authors: Huub van Dam, Manhui Wang, Andrew G. Sunderland, Ian J. Bush, Peter J. Knowles, Martyn Guest

Quantum chemistry calculations have a compelling requirement for one-sided communication given their irregular data access patterns and irregular task sizes. Hence the MPI-2 passive target one-sided communications seem attractive as a standards based foundation for the implementation of the main algorithms in the field. For successful deployment it is important that good performance can be achieved reliably across a wide variety of platforms. We have tested this performance using a simple shared counter kernel across a number of machines and MPI implementations, including both open source MPI implementations - OpenMPI and MPICH2 - and those on the current and previous generation of the UK’s national academic supercomputers, HECToR and HPCx. We find that the performance varies greatly with the MPI implementation in question. The speed with which communications are progressed was found to vary up to 4 orders of magnitude in the kernel program depending on the MPI library used. As a result the "time to solution" for the kernel could vary by as much as about a factor 2, although greater impacts are anticipated for more complex algorithms such as Fock-builders. This spread in performance relates to two aspects of the MPI-2 standard. The first is whether the MPI implementation takes a minimalistic approach to satisfy rule 11.7.2 on the progress of one-sided communications or goes beyond that. Secondly, does the MPI library implement full multi-threading support, as specified in section 12.4, or not. Overall, we find that the MPI-2 standard is not strong enough to guarantee that MPI libraries provide acceptable performance characteristics to use one-sided communications effectively. We suggest that the standard be strengthened to address these performance issues – without this it seems unlikely that the MPI-2 one-sided communications will be useful for real world applications.

Ab Initio potential grid based docking: From High Performance Computing to In Silico Screening

AIP Conference Proceedings 940   October 4, 2007

Authors: Huub van Dam, Marc R. de Jonge, H. Maarten Vinkers, Joop Van Lenthe, Frits Daeyaert, Ian J. Bush, Paul Sherwood, Martyn Guest

We present a new method for the generation of potential grids for protein-ligand docking. The potential of the docking target structure is obtained directly from the electron density derived through an ab initio computation. A large subregion was selected to allow the full ab initio treatment of a the Isocitrate Lyase enzyme. The electrostatic potential is tested by docking a small charged molecule (succinate) into the binding site. The ab initio grid yields a superior result by producing the best binding orientation and position, and by recognizing it as the best. In contrast the same docking procedure, but using a classical point-charge based potential, produces a number of additional incorrect binding poses, and does not recognize the correct pose as the best solution.

Point defects in ZnO

Faraday Discussions   2006

Authors: Huub van Dam, Alexey A. Sokol, Samuel A. French, Stefan T. Bromley, C. Richard A. Catlow, Paul Sherwood

We have investigated intrinsic point defects in ZnO and extended this study to Li, Cu and Al impurity centres. Atomic and electronic structures as well as defect energies have been obtained for the main oxidation states of all defects using our embedded cluster hybrid quantum mechanical/molecular mechanical approach to the treatment of localised states in ionic solids. With these calculations we were able to explain the nature of a number of experimentally observed phenomena. We show that in zinc excess materials the energetics of zinc interstitial are very similar to those for oxygen vacancy formation. Our results also suggest assignments for a number of bands observed in photoluminescence and other spectroscopic studies of the material.

Formation of Heteroatom Active Sites in Zeolites by Hydrolysis and Inversion

Angewandte Chemie   January 30, 2006

Authors: Huub van Dam, Judy To, Alexey A. Sokol, Samuel A. French, C. Richard A. Catlow, Paul Sherwood

Quantenmechanische/molekülmechanische Rechnungen wurden verwendet, um die Strukturen der aktiven Zentren in Silicalit zu untersuchen, der mit Heteroatomen der Gruppen 4 und 14 dotiert ist. Nach diesen

Rechnungen sind tripodale (siehe Bild; Si gelb, O rot, H weiß, Heteroatom blau) oder bipodale Zentren, die durch Hydrolyse und Inversion entstehen, bevorzugt.

QM/MM Modelling of the TS-1 Catalyst using HPCx

Journal of Materials Chemistry 2006

Authors: Huub van Dam, Judy To, Paul Sherwood, Alexey A. Sokol, Ian J. Bush, C. Richard A. Catlow, Samuel A. French, Martyn Guest

We report a series of computations on the active site in Ti-substituted zeolites, specifically TS-1. Hybrid QM/MM methods based on density functional calculations using the BB1K functional and a valence force field are used to study the processes of hydrolysis of Ti–O–Si linkages and inversion of the TiO_4 tetrahedra. The structural features of the resulting series of tetra- and tripodal Ti moieties are in good agreement with data from EXAFS studies. The suggestion is made that the tripodal species will dominate in hydrous conditions, and that this is likely to be the chemically active form. We have made extensive use of the massively parallel HPCx computer system for these investigations and outline some of the technical developments to the ChemShell software that were needed to support the study.

Large Scale Electronic Structure Calculations in the Study of the Condensed Phase

Journal of Molecular Structure: Theochem October 12, 2006

Authors: Huub van Dam, Martyn Guest, Paul Sherwood, Jens M. H. Thomas, Joop Van Lenthe, Joost van Lingen, Christine L. Bailey, Ian J. Bush

We consider the role that large-scale electronic structure computations can now play in the modelling of the condensed phase. To structure our analysis, we consider four distinct ways in which today's scientific targets can be re-scoped to take advantage of advances in computing resources:

1. Time to solution—performing the same calculation, with delivery of the simulation in shorter elapsed time;

2. Size—applying today's methods to a more extensive problem;

3. Accuracy—replacing current physical models with more accurate ones.

4. Sampling—simultaneously studying more chemical or conformational states.

Each of these offer some scientific rewards, but all present technological challenges and it is likely that a mixture of approaches will be needed to make the best use of capability computing. We discuss some aspects of our work in each of these areas, including replicated and distributed data parallel implementations of GAMESS-UK, and approaches incorporating multi-level parallelism. Examples of the latter include pathway optimization using replica methods and task farming approaches to global optimization problems.

We consider a typical application of GAMESS-UK to heterogeneous catalysis, and the role that large-scale DFT vibrational frequency calculations have played in the study of partial oxidation catalysts incorporating supported VO_x species.

Starting SCF calculations by superposition of atomic densities

Journal of Computational Chemistry   March 23, 2006

Authors: Huub van Dam, Joop Van Lenthe, Renate Zwaans, Martyn Guest

We describe the procedure to start an SCF calculation of the general type from a sum of atomic electron densities, as implemented in GAMESS-UK. Although the procedure is well known for closed-shell calculations and was already suggested when the Direct SCF procedure was proposed, the general procedure is less obvious. For instance, there is no need to converge the corresponding closed-shell Hartree-Fock calculation when dealing with an open-shell species. We describe the various choices and illustrate them with test calculations, showing that the procedure is easier, and on average better, than starting from a converged minimal basis calculation and much better than using a bare nucleus Hamiltonian.

The GAMESS-UK electronic structure package: algorithms, developments and applications

Molecular Physics   March 2005

Authors: Huub van Dam, Martyn Guest, Ian J. Bush, Paul Sherwood, Jens M. H. Thomas, Joop Van Lenthe, Remco Havenith, John Kendrick

A description of the ab initio quantum chemistry package GAMESS-UK is presented. The package offers a wide range of quantum mechanical wavefunctions, capable of treating systems ranging from closed-shell molecules through to the species involved in complex reaction mechanisms. The availability of a wide variety of correlation methods provides the necessary functionality to tackle a number of chemically important tasks, ranging from geometry optimization and transition-state location to the treatment of solvation effects and the prediction of excited state spectra. With the availability of relativistic ECPs and the development of ZORA, such calculations may be performed on the entire Periodic Table, including the lanthanides. Emphasis is given to the DFT module, which has been extensively developed in recent years, and a number of other, novel features of the program. The parallelization strategy used in the program is outlined, and detailed speedup results are given. Applications of the code in the areas of enzyme and zeolite catalysis and in spectroscopy are described.

Large scale computing with GAMESS-UK (COMP 123)

Abstracts of Papers of the American Chemical Society   2005

Authors: Huub van Dam, Joop Van Lenthe, Martyn Guest, Ian J. Bush, Paul Sherwood, Jens M. H. Thomas, Remco Havenith, Joost van Lingen, Marc de Jonge

The quantumchemical program GAMESS-UK has known a long history of parallelisation. As machines with more than 500 processors and gigabytes of memory have become available, it has become relevant to establish the current limits of the program and the scaling behaviour. We will discuss some of the recent developments in GAMESS-UK, which help exploiting the current capability machines. We will discuss the approaches utilised in parallellising the HF/DFT, CI and Valence Bond modules. We will demonstrate the ability to perform HF/DFT calculations on large (bio)molecules, with more than 3000 atoms and 11000 basis functions, in calculations of molecular potential and densities, meant as input for Computer Aided Drug Design endeavors. We will show DFT Hessian calculations on molecules with more than 400 basis functions and over 50 atoms which are being used to elucidate the structure of heterogenous Vanadium catalysts, by analysis of their infrared and Raman spectra.

Minimal dipole charge for a dipole-bound dianion

Molecular Physics   August 2003

Authors: Huub van Dam, Inaki Silanes, Jesus M. Ugalde

Full configuration interaction calculations for two electrons moving in the field of a fixed finite dipole (FFD) have been carried out, in order not only to determine the conditions for stability relative to one electron detachment, but also to check whether a true dipole-bound system, converse to a Stark-shifted atom-like system, is generated. The FFD model is constructed by placing two point charges of absolute value q, and opposite sign, separated by a distance d.

It has been found that, although dipole charges as small as q#0.91 au can bind two electrons stronger than one for large enough values of d (thus giving stable systems relative to the detachment of one electron), it is only for q > 2.941 au that the length of the dipole is short enough for the size of the electronic cloud to actually exceed it, so that the system can really be regarded as a dipole-bound dianion, and not a Stark-shifted atom.

Theoretical exploration of ultrafast spectroscopy of small clusters

International Journal of Quantum Chemistry   June 21, 2001

Authors: Huub van Dam, Vlasta Bona#i#-Koutecký, Michael Hartmann, Ji#í Pittner

The central issue in femtosecond (fs) time resolved spectroscopy of clusters is the investigation of geometric relaxation and internal vibrational redistribution (IVR) after optical excitation in a nonequilibrium configuration of nuclei by laser photoelectron excitation, and corresponding time delayed probing by multiphoton-ionization. For this purpose, we have developed multistate ab initio molecular dynamics involving adiabatic ground and excited electronic states, as well as nonadiabatic coupling between them, using the time evolution of initial thermal ensemble in Wigner representation. The combination of ab initio quantum chemical methods, used for the adiabatic and nonadiabatic molecular dynamics “on the fly,” and the Wigner distribution approach for the description of the motion of the nuclei allowed us the accurate determination of pump-probe and pump-dump signals also under temperature dependent initial conditions...

2nd derivatives of the electronic energy in density functional theory

DL Technical Reports   July 26, 2001

Authors: Huub van Dam

This document details the equations needed to implement the calculation of vibrational frequencies within the density functional formalism of electronic structure theory. This functionality has been incorporated into the CCP1 DFT module and the required changes to the application programmers interface are outlined. Throughout it is assumed that an implementation of Hartree-Fock vibrational frequencies is available that can be modified to incorporate the density functional formalism. Employing GAMESS-UK as an example the required changes to the Hartree-Fock code are outlined.

Theoretical exploration of stationary and of ultrafast spectroscopy of small clusters

Applied Physics B: Lasers and Optics   2000

Authors: Huub van Dam, Michael Hartmann, Jiri Pittner, Vincent Veyret, Vlasta Bonacic-Koutecky

Stationary spectra offer information on the interplay between the structures and the nature of electronic excitations reflecting bonding properties, as shown by comparing Si_n with Ag_n (n = 4#6) clusters. In order to study the dynamical properties, simulations and analysis of femtosecond (fs) time-resolved pump–probe or pump–dump signals have been carried out, which allows us to determine the timescales and the nature of configurational changes versus internal vibrational relaxation (IVR) in electronic ground or excited states. For this purpose we have developed a multi-state ab initio molecular dynamics (involving ground as well as adiabatic or non-adiabatic excited electronic states) on the timescale of the nuclear motion, using the time evolution of a thermal ensemble in the Wigner representation.

The combination of ab initio quantum-chemical methods used for the molecular dynamics ‘on the fly’ and the Wigner-distribution approach for the description of the motion of the nuclei also allowed the accurate determination of pump–probe and pump–dump signals under temperature dependent initial conditions. We use this novel combination of methods to investigate the dynamics in excited states of non-stoichiometric Na_nF_n#1 clusters with a single excess electron. The timescales of the structural relaxation in excited states versus intramolecular vibrational relaxation processes have been determined, as illustrated for the example of Na_4F_3. This is the first study of the system with 15 degrees of freedom for which the dynamics in the excited states has been carried out without the precalculation of the energy surfaces.

Algorithms, developments and applications in molecular modelling: The GAMESS-UK ab initio code

AIP Conference Proceedings 479   December 9, 1998

Authors: Huub van Dam, Martyn Guest

Focusing on molecular electronic structure, an outline is presented of the range of methods commonly used in computational chemistry, with consideration given to the accuracy, limitations and performance costs of the HF, DFT, MP2 and CCSD methods in understanding neutron-scattering related phenomena. Using GAMESS-UK as an example of a large electronic structure package, we illustrate the capabilities of these methods in computing the geometrical structures of transition metal complexes. These studies both reinforce the value of DFT, and reveal an overall inconsistency in the MP2-based predictions. Finally, emphasis is given to the cost effectiveness and applicability of the spectrum of hardware available to the computational chemist, from the single-user PC to massively parallel processors (MPP). We illustrate the potential of the latter platforms in enhancing the level of desk-top simulations by two-orders of magnitude, and demonstrate the potential of parallel PC-based Beowulf Systems.

Theoretical study of femtosecond pump-probe signals of nonstoichiometric alkali halide clusters

The European Physical Journal D - Atomic, Molecular, Optical and Plasma Physics   November 22, 1998

Authors: Huub van Dam, Michael Hartmann, Jiri Pittner, Vlasta Bonacic-Koutecky

We present the outline of the ab-initio molecular dynamics method for adiabatic excited states which is suitable to treat the involved excited electronic states simultaneously with the motion of the nuclei. The method allows the simulation of time-resolved pump-probe signals and the analysis of the signals in terms of the underlying nuclear dynamics. Moreover, the timescales of the structural relaxation in excited electronic states vs. intramolecular vibrational relaxation processes can be estimated and the influence of the initial temperature of the cluster can be determined.

The method is applied to the analysis of the timescales as well as of the dynamics of localized/delocalized electronic excitations in the Na_2F molecule which represents a test example for nonstoichiometric alkali-halides with one excess electron. The aim is to develop a method for the investigation of time-dependent laser induced processes in clusters which is capable to extract the leading physical effects responsible for the controlled excitation dynamics and for the evolution of the non-equilibrium states.

The excitation energies of 1,1'-bicyclohexylidene and 1,1':4',1"-tercyclohexylidene. A comparison of multi-reference configuration interaction and perturbation theory approaches

Chemical Physics   July 19, 1999

Authors: Huub van Dam, Remco Havenith, Joop Van Lenthe, Leonardus W. Jenneskens

The lowest valence transition energies of the first two homologues of the oligo(cyclohexylidene) series, viz. 1,1#-bicyclohexylidene (1) and 1,1#:4#,1#-tercyclohexylidene (2), were computed using multi-reference perturbation theory at the second (MR-MP2) and third order (MR-MP3). The results were compared to the previously reported results using a non-selected multi-reference single-double CI (Direct-CI) and a selected multi-reference single-double CI (MRDCI) approach, in order to obtain insight in the applicability of the MRDCI approach for the prediction of absorption spectra of large organic molecules. Calculations of the lowest valence excited states of 1,1#-bicyclohexylidene (1) showed that MR-MP3 gives energies comparable to those of Direct-CI; for the transition energies, MR-MP2 and MR-MP3 performed equally well. The calculations on the excited states of 1,1#:4#,1#-tercyclohexylidene (2) revealed that the MR-MP2 methodology is not reliable for the prediction of its transition energies, if the multi-reference character of the excited states increases. Although remarkable differences in the absolute transition energies were found, the assignments of the lowest valence transitions of 1,1#-bicyclohexylidene (1) and 1,1#:4#,1#-tercyclohexylidene (2) were similar at all levels of theory.

Exact size consistency of multi-reference Møller-Plesset perturbation theory

International Journal of Quantum Chemistry   March 3, 1999

Authors: Huub van Dam, Joop Van Lenthe, Paul J. A. Ruttink

Single-reference closed-shell Møller–Plesset perturbation theory is well known for its size consistency, a quality that is essential for consistent comparisons of calculations on molecules of different size. However, it is far from obvious whether this quality can be retained in the multireference case. In this work it is shown that an exactly size consistently generalization to multireference perturbation theory can be constructed. The central result is that the zeroth-order Hamiltonian should be constructed using separate projection operators for each excitation level, i.e., it should contain no couplings between different excitation levels.

The size-consistency of multi-reference Møller-Plesset perturbation theory

Molecular Physics   February 1998

Authors: Huub van Dam, Joop Van Lenthe, Peter Pulay

The size consistency of multi-reference Møller-Plesset perturbation theory as a function of the structure of the zeroth-order Hamiltonian is studied. In calculations it is shown that the choice of projection operators to define the zeroth-order Hamiltonian is crucial. In essence whenever such a projection operator can be written as the sum of projection operators onto particular subspaces, cross-product terms may appear in the zeroth-order Hamiltonian that spoil the size consistency. This problem may be solved using a separate projection operator for each subspace spanning an excitation level. In principle a zeroth-order Hamiltonian based on these projection operators results in a size consistent perturbation theory. However, it was found that some non-local spin recoupling effects remain. A new zeroth-order Hamiltonian formulated recently circumvents this problem and is shown to be exactly size consistent. Apart from the choice of projection operators, the orthogonalization of the excited states is crucial also. It was found that modified Gramm-Schmidt in quadruple precision was not sufficient. A pivotted Householder QR factorization (in double precision) offered the numerical stability needed to obtain size consistent results.

Calculations on correlation effects in molecules: Convergence and size-consistency of multi-reference methods

PhD Thesis, Utrecht University   October 13, 1997

Authors: Huub van Dam

This thesis is about methods for electronic structure calculations on molecular systems. The ultimate goal is to construct methods that yield potential energy surfaces of sufficient accuracy to allow a qualitatively correct description of the chemistry of these systems; i.e. heat of formation, isomerisation barriers, equilibrium geometries, and vibrational spectra. In order to properly calculate the potential energy surfaces for all these properties a multi-configurational starting point is essential. This means that all methods that will be discussed are based on a multi-reference wavefunction where the reference function is optimised using the multi-configurational Hartree-Fock (MCHF) method...

Convergence behaviour of multi-reference perturbation theory: An indicator

Molecular Physics   April 1997

Authors: Huub van Dam, Joop Van Lenthe

An approach to investigate the convergence behaviour of a (multireference) Moeller-Plesset perturbation series is proposed in which the series is evaluated up to arbitrary order without reference to excitations beyond doubles. It is shown that this approach qualitatively reproduces the convergence behaviour of the complete perturbation series for stretched H_2 molecules and H_2O molecules. The MRMP approach does show improved convergence characteristics compared with single reference MP for stretched H_2 molecules. Applications of the approach to O_2 at 4·0 a_0, and N_2 at 5·0 a_0 show that the MRMP series is divergent for these cases. Furthermore, in the latter case it was found that the multireference series shows much stronger divergence than the single reference series.

An improvement of Davidsons iteration method: Applications the MRCI and MRCEPA(0) calculations

Journal of Computational Chemistry 17, 267-272   1996

Authors: Huub van Dam, Joop Van Lenthe, Gerard L. G. Sleijpen, Henk A. van der Vorst

Davidson's method is widely used for finding the lowest eigenvalues of large matrices. Recently, mathematicians have shown that Davidson's derivation could be improved. They have corrected the derivation yielding a new iteration method. In this article this new method is adapted for realistic MRCI and MRCEPA calculations. Results show that the new method converges significantly faster in H_2O and O_2 with moderately elongated bonds than Davidson's original method. The new method offers new insights into the rate of convergence of Davidson's original method.

NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations

Computer Physics Communications   September 2010

Authors: Huub van Dam, Marat Valiev, Eric J. Bylaska, Niri Govind, Karol Kowalski, Tjerk P. Straatsma, Dunyou Wang, Jarek Nieplocha, Edo Apra, Theresa L. Windus, Bert de Jong

The latest release of NWChem delivers an open-source computational chemistry package with extensive capabilities for large scale simulations of chemical and biological systems. Utilizing a common computational framework, diverse theoretical descriptions can be used to provide the best solution for a given scientific problem. Scalable parallel implementations and modular software design enable efficient utilization of current computational architectures. This paper provides an overview of NWChem focusing primarily on the core theoretical modules provided by the code and their parallel performance.

Utilizing high performance computing for chemistry: parallel computational chemistry

Physical Chemistry and Chemical Physics 2010

Authors: Huub van Dam, Bert de Jong, Eric J. Bylaska, Niri Govind, Curtis L. Janssen, Karol Kowalski, Thomas Müller, Ida M. B. Nielsen, Valera Veryazov, Roland Lindh

Parallel hardware has become readily available to the computational chemistry research community. This perspective will review the current state of parallel computational chemistry software utilizing high-performance parallel computing platforms. Hardware and software trends and their effect on quantum chemistry methodologies, algorithms, and software development will also be discussed.

Fault Tolerant Communication Runtime Support for Data Centric Programming Models

HiPC 2010, Goa, India December 2010

Authors: Huub van Dam, Abhinav Vishnu, Bert de Jong, Pavan Balaji, Song Leon Shuaiwen

The largest supercomputers in the world today consist of hundreds of thousands of processing cores and many more other hardware components. At such scales, hardware faults are a commonplace, necessitating fault-resilient software systems. While different fault-resilient models are available, most focus on allowing the computational processes to survive faults. On the other hand, we have recently started investigating fault resilience techniques for data-centric programming models such as the partitioned global address space (PGAS) models. The primary difference in data-centric models is the decoupling of computation and data locality. That is, data placement is decoupled from the executing processes, allowing us to view process failure (a physical node hosting a process is dead) separately from data failure (a physical node hosting data is dead).

In this paper, we take a first step toward data-centric fault resilience by designing and implementing a fault0resilient, onesided communication runtime framework using Global Arrays and its communication system, ARMCI. The framework consists of a fault-resilient process manager; low-overhead and networkassisted remote-node fault detection module; non-data-moving collective communication primitives; and failure semantics and error codes for one-sided communication runtime systems. Our performance evaluation indicates that the framework incurs little overhead compared to state-of-the-art designs and provides a fundamental framework of fault resiliency for PGAS models.

Designing a Scalable Fault Tolerance Model for High Performance Computational Chemistry: A Case Study with Coupled Cluster Perturbative Triples

Journal of Chemical Theory and Computation   December 8, 2010

Authors: Huub van Dam, Abhinav Vishnu, Bert de Jong

In the past couple of decades, the massive computational power provided by the most modern supercomputers has resulted in simulation of higher-order computational chemistry methods, previously considered intractable. As the system sizes continue to increase, the computational chemistry domain continues to escalate this trend using parallel computing with programming models such as Message Passing Interface (MPI) and Partitioned Global Address Space (PGAS) programming models such as Global Arrays. The ever increasing scale of these supercomputers comes at a cost of reduced Mean Time Between Failures (MTBF), currently on the order of days and projected to be on the order of hours for upcoming extreme scale systems. While traditional disk-based check pointing methods are ubiquitous for storing intermediate solutions, they suffer from high overhead of writing and recovering from checkpoints. In practice, checkpointing itself often brings the system down. Clearly, methods beyond checkpointing are imperative to handling the aggravating issue of reducing MTBF. In this paper, we address this challenge by designing and implementing an efficient fault tolerant version of the Coupled Cluster (CC) method with NWChem, using in-memory data redundancy. We present the challenges associated with our design, including an efficient data storage model, maintenance of at least one consistent data copy, and the recovery process. Our performance evaluation without faults shows that the current design exhibits a small overhead. In the presence of a simulated fault, the proposed design incurs negligible overhead in comparison to the state of the art implementation without faults.

SO2-Binding Properties of Cationic #6,#1-NCN-Pincer Arene Ruthenium Platinum Complexes: Spectroscopic and Theoretical Studies

Dalton Transactions   February 4, 2011

Authors: Huub van Dam, Sylvestre Bonnet, Joop Van Lenthe, Gerard van Koten, Bert Klein Gebbink

The SO2-binding properties of a series of #6,#1-NCN-pincer ruthenium platinum complexes have been studied by both UV-visible spectroscopy, and theoretical calculations. When an electronwithdrawing [Ru(C5R5)]+ fragment (R = H or Me) is #6-coordinated to the phenyl ring of the NCNpincer platinum fragment (cf. [2]+ and [3]+, see scheme 1), the characteristic orange coloration (pointing to #1- SO2 binding to Pt) of a solution of the parent NCN-pincer platinum complex 1 in dichloromethane upon SO2-bubbling is not observed. However, when the ruthenium center is #6- coordinated to a phenyl substituent linked in para-position to the carbon-to-platinum bond, i.e. complex [4]+, the SO2-binding property of the NCN-platinum center seems to be retained, as bubbling SO2 into a solution of the latter complex produces the characteristic orange color. We performed theoretical calculations at the MP2 level of approximation and TD-DFT studies, which enabled us to interpret the absence of color change in the case of [2]+ as an absence of coordination of SO2 to platinum. We analyze this absence or weaker SO2-coordination in dichloromethane to be a consequence of the relative electron-poorness of the platinum center in the respective #6- ruthenium coordinated NCN-pincer platinum complexes, that leads to a lower binding energy and an elongated calculated Pt-S bond distance. We also discuss the effects of electrostatic interactions in these cationic systems, which also seems to play a destabilizing role for complex [2(SO2)]+.

MSC Benchmark Revision 1.0

Pacific Northwest National Laboratories   January 2011

Authors: Huub van Dam, Kurt Glaesemann, Jared Carr

The MSC Benchmark Revision 1.0 is designed to match the expected utilization of MSC computer resources. The PDF document describes a set of benchmarks, which reflect the diverse workload the MSC capability enables for EMSL users. For more information, please see the documentation and run rules.

The benchmarks are based upon publicly available codes from other groups and local Pacific Northwest National Laboratory (PNNL) codes (a README.pnnl file is included with each code to provide suggestions for building the benchmark).

PNNL document number: PNNL-SA-77159.

NWChem: Scalable parallel computational chemistry

Interdisciplinary Reviews: Computational Molecular Science   May 6, 2011

Authors: Huub van Dam, Bert de Jong, Eric Bylaska, Niri Govind, Karol Kowalski, Tjerk Straatsma, Marat Valiev

NWChem is a general-purpose computational chemistry code specifically designed to run on distributed memory parallel computers. The core functionality of the code focuses on molecular dynamics, Hartree-Fock theory, and density functional theory methods for both plane-wave basis sets as well as Gaussian basis sets, tensor contraction engine-based coupled cluster capabilities, and combined quantum mechanics/molecular mechanics descriptions. It was realized from the beginning that scalable implementations of these methods required a programming paradigm inherently different from what message-passing approaches could offer. In response, a global address space library, the Global Array toolkit, was developed. The programming model it offers is based on using predominantly one-sided communication. This model underpins most of the functionality in NWChem, and the power of it is exemplified by the fact that the code scales to tens of thousands of processors. In this paper, the core capabilities of NWChem are described as well as their implementation to achieve an efficient computational chemistry code with high parallel scalability.

Massively parallel implementation of the multireference Brillouin-Wigner CCSD method

Chemical Physics Letters   August 22, 2011

Authors: Huub van Dam, Jiri Brabec, Sriram Krishnamoorthy, Karol Kowalski, Jiri Pittner

This Letter reports the parallel implementation of the Multireference Brillouin–Wigner Coupled Cluster method with Single and Double excitations (MR BWCCSD). Preliminary tests for systems composed of 304 and 440 correlated orbitals demonstrate the performance of our implementation across 1000 cores and clearly indicate the advantages of using improved task scheduling. Possible ways for further improvements of the parallel performance are also delineated.

Parallel Quantum Chemistry at the Crossroads

IGI-global   October 2011

Authors: Huub van Dam

Quantum chemistry was a compute intensive field from the beginning. It was also an early adopter of parallel computing, and hence, has more than twenty years of experience with parallelism. However, recently parallel computing has seen dramatic changes, such as the rise of multi-core architectures, hybrid computing, and the prospect of exa-scale machines requiring 1 billion concurrent threads. It is doubtful that current approaches can address the challenges ahead. As a result, the field finds itself at a crossroads, facing the challenge to successfully identify the way forward. This chapter tells a story in two parts. First, the achievements to date are considered, offering insights learned so far. Second, we look at paradigms based on directed acyclic graphs (DAG). The computer science community is strongly advocating this paradigm, but the quantum chemistry community has no experience with this approach. Therefore recent developments in that area will be discussed and their suitability for future parallel quantum chemistry computing demands considered.

Providing fault tolerance in extreme scale parallel applications: What can the HPC community learn from the Database community?

HPCDB   November 18, 2011

Authors: Huub van Dam, Abhinav Vishnu, Bert de Jong

The database community has a long history of considering fault tolerance, in HPC fault tolerance has been limited to simple checkpoint/restart strategies. On emerging extreme scale compute facilities these simple techniques are no longer feasible and the HPC community will need to consider more advanced strategies.

In this paper we discuss ongoing and planned work regarding fault tolerance in the context of large scale HPC calculations. Issues ranging from the detection of hard faults, recovery and continued execution, to handling soft errors will be discussed. As fault tolerance is a cross cutting theme for both databases and HPC applications it is hoped that this position paper will lead to a lively exchange of insights.

Universal state-selective corrections to multi-reference coupled-cluster theories with single and double excitations

Journal of Chemical Physics   March 22, 2012

Authors: Huub van Dam, Jiri Brabec, Karol Kowalski, Jiri Pittner

The recently proposed universal state-selective (USS) corrections [K. Kowalski, J. Chem. Phys. 134, 194107 (2011)] to approximate multi-reference coupled-cluster (MRCC) energies can be commonly applied to any type of MRCC theory based on the Jeziorski-Monkhorst [B. Jeziorski and H. J. Monkhorst, Phys. Rev. A 24, 1668 (1981)] exponential ansatz. In this paper we report on the performance of a simple USS correction to the Brillouin-Wigner and Mukherjee's MRCC approaches employing single and double excitations (USS-BW-MRCCSD and USS-Mk-MRCCSD). It is shown that the USS-BW-MRCCSD correction, which employs the manifold of single and double excitations, can be related to a posteriori corrections utilized in routine BW-MRCCSD calculations. In several benchmark calculations we compare the USS-BW-MRCCSD and USS-Mk-MRCCSD results with the results obtained with the full configuration interaction method.

Parallel Implementation of Multireference Coupled-Cluster Theories Based on the Reference-Level Parallelism

Journal of Chemical Theory and Computation   January 11, 2012

Authors: Huub van Dam, Jiri Brabec, Jiri Pittner, Edoardo Aprà, Karol Kowalski

A novel algorithm for implementing a general type of multireference coupled-cluster (MRCC) theory based on the Jeziorski–Monkhorst exponential ansatz [Jeziorski, B.; Monkhorst, H. J. Phys. Rev. A1981, 24, 1668] is introduced. The proposed algorithm utilizes processor groups to calculate the equations for the MRCC amplitudes. In the basic formulation, each processor group constructs the equations related to a specific subset of references. By flexible choice of processor groups and subset of reference-specific sufficiency conditions designated to a given group, one can ensure optimum utilization of available computing resources. The performance of this algorithm is illustrated on the examples of the Brillouin–Wigner and Mukherjee MRCC methods with singles and doubles (BW-MRCCSD and Mk-MRCCSD). A significant improvement in scalability and in reduction of time to solution is reported with respect to recently reported parallel implementation of the BW-MRCCSD formalism [Brabec, J.; van Dam, H. J. J.; Kowalski, K.; Pittner, J. Chem. Phys. Lett.2011, 514, 347].

Towards large-scale calculations with state-specific multireference coupled cluster methods: Studies on dodecane, naphthynes, and polycarbenes

Chemical Physics Letters June 4, 2012

Authors: Huub van Dam, Jiri Brabec, Kiran Bhaskaran Nair, Jiri Pittner, Karol Kowalski

Multireference Coupled Cluster (MRCC) methods are essential for accurate descriptions of molecular systems with degenerate states. However, their use has been hampered by their high computational cost, which can be alleviated with scalable MRCC codes. Exploiting large computational resources our recent parallelization work has demonstrated MRCC on chemically relevant molecules within practical per core resource constraints.

We report on parallel implementations of Mukherjee’s MRCC (MR MkCCSD) and Brillouin-Wigner MRCC (MR BWCC) models. Especially abstracting a methyl-group from dodecane, and singlet-high-spin splittings in naphthyne isomers and polycarbenes demonstrate systems with significant multireference character and size that currently can be handled.

Implementation of the multireference Brillouin-Wigner and Mukherjee's coupled cluster methods with non-iterative triple excitations utilizing reference-level parallelism

Journal of Chemical Physics September 7, 2012

Authors: Huub van Dam, Kiran Bhaskaran-Nair, Jiri Brabec, Edoardo Aprà, Jiri Pittner, Karol Kowalski

In this paper we discuss the performance of the non-iterative state-specific multireference coupled cluster (SS-MRCC) methods accounting for the effect of triply excited cluster amplitudes. The corrections to the Brillouin-Wigner and Mukherjee's MRCC models based on the manifold of singly and doubly excited cluster amplitudes (BW-MRCCSD and Mk-MRCCSD, respectively) are tested and compared with exact full configuration interaction results for small systems (H2O, N2, and Be3). For the larger systems (naphthyne isomers) the BW-MRCC and Mk-MRCC methods with iterative singles, doubles, and non-iterative triples

(BW-MRCCSD(T) and Mk-MRCCSD(T)) are compared against the results obtained with single reference coupled cluster methods. We also report on the parallel performance of the non-iterative implementations based on the use of processor groups.

Noniterative multireference coupled cluster methods on heterogeneous CPU-GPU systems

Journal of Chemical Theory and Computation   March 6, 2013

Authors: Huub van Dam, Kiran Bhaskaran-Nair, Wenjing Ma, Sriram Krishnamoorthy, Oreste Villa, Edoardo Aprà, Karol Kowalski

A novel parallel algorithm for noniterative multireference coupled cluster (MRCC) theories, which merges recently introduced reference-level parallelism (RLP) [Bhaskaran-Nair, K.; Brabec, J.; Aprà, E.; van Dam, H. J. J.; Pittner, J.; Kowalski, K. J. Chem. Phys.2012, 137, 094112] with the possibility of accelerating numerical calculations using graphics processing units (GPUs) is presented. We discuss the performance of this approach applied to the MRCCSD(T) method (iterative singles and doubles and perturbative triples), where the corrections due to triples are added to the diagonal elements of the MRCCSD effective Hamiltonian matrix. The performance of the combined RLP/GPU algorithm is illustrated on the example of the Brillouin–Wigner (BW) and Mukherjee (Mk) state-specific MRCCSD(T) formulations.

A case for soft error detection and correction in computational chemistry

Journal of Chemical Theory and Computation   July 19, 2013

Authors: Huub van Dam, Abhinav Vishnu, Bert de Jong

High performance computing platforms are expected to deliver 1.e18 floating operations per second by the year 2022 through the deployment of millions of cores. Even if every core is highly reliable the sheer number of them will mean that the mean time between failures will become so short that most applications runs will suffer at least one fault. In particular soft errors caused by intermittent incorrect behavior of the hardware are a concern as they lead to silent data corruption. In this paper we investigate the impact of soft errors on optimization algorithms using Hartree-Fock as a particular example. Optimization algorithms iteratively reduce the error in the initial guess to reach the intended solution. Therefore they may intuitively appear to be resilient to soft errors. Our results show that this is true for soft error s of small magnitudes but not for large errors. We suggest error detection and correction mechanisms for different classes of data structures. The results obtained with these mechanisms indicate that we can correct more than 95% of the soft errors at moderate increases in the computational cost.

Graphical user environments for scientific computing

UKHEC Technical Report   2003

Authors: Mike Ashworth, Rob Allan, Chris Mueller, Huub van Dam, Bill Smith, Daniel Hanlon, Barry Searle, Andrew Sunderland

Many scientific applications now have a Graphical User Interface (GUI) customised to make its use more intuitive for novices and experts alike. It is interesting to compare technology designed for GUI development and evaluate the kind of functionality commonly incorporated.

GUIs take on a particular importance in the computational Grid world where aspects of managing applications in a heterogeneous distributed environment can be hidden from most users. Indeed projects like UNICORE provide a GUI specifically designed for submitting HPC applications transparently to a range of different systems. Other similar projects provide seamless access to data across a range of file servers or databases. Our examples are, in the main, taken from work currently ongoing at Daresbury Laboratory in support of the CCP, HPCI and e-Science programmes. We describe implementations in Java, Perl/Tk, Python/Tk and C++/AVS Express and also C and Perl/CGI. There is particular emphasis on using Web browsers as a special form of GUI because of the current focus on the computational Grid and Internet based distributed services (Web services).

Simulating one-photon absorption and resonance Raman scattering spectra using analytical excited state energy gradients within time-dependent density functional theory

Journal of Chemical Theory and Computation   November 6, 2013

Authors: Daniel Silverstein, Niri Govind, Huub van Dam, Lasse Jensen

A parallel implementation of analytical time-dependent density functional theory gradients is presented for the quantum chemistry program NWChem. The implementation is based on the Lagrangian approach developed by Furche and Ahlrichs. To validate our implementation, we first calculate the Stokes shifts for a range of organic dye molecules using a diverse set of exchange-correlation functionals (traditional density functionals, global hybrids, and range-separated hybrids) followed by simulations of the one-photon absorption and resonance Raman scattering spectrum of the phenoxyl radical, the well-studied dye molecule rhodamine 6G, and a molecular host–guest complex (TTF_C BPQT^4+). The study of organic dye molecules illustrates that B3LYP and CAM-B3LYP generally give the best agreement with experimentally determined Stokes shifts unless the excited state is a charge transfer state. Absorption, resonance Raman, and fluorescence simulations for the phenoxyl radical indicate that explicit solvation may be required for accurate characterization. For the host–guest complex and rhodamine 6G, it is demonstrated that absorption spectra can be simulated in good agreement with experimental data for most exchange-correlation functionals. However, because one-photon absorption spectra generally lack well-resolved vibrational features, resonance Raman simulations are necessary to evaluate the accuracy of the exchange-correlation functional for describing a potential energy surface.

Describing Excited State Relaxation and Localization in TiO 2 Nanoparticles Using TD-DFT

Journal of Chemical Theory and Computation   December 2014

Authors: Enrico Berardo, Hanshi Hu, Huub van Dam, Stephen Shevlin, Scott Woodley, Kowalski Karol, Martijn Zwijnenburg

Building Scalable PGAS Communication Subsystem on Blue Gene/Q

2013 IEEE International Symposium on Parallel & Distributed Processing   May 2013

Authors: Abhinav Vishnu, Darren Kerbyson, Kevin Barker, Huub van Dam

Diagnosing the causes and severity of one-sided message contention

Proceedings of the 20th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming   2015

Authors: Nathan Tallent, Abhinav Vishnu, Huub van Dam, Jeff Daily, Darren Kerbyson, Adolfy Hoisie

Generalization of the Kohn-Sham system that can represent arbitrary one-electron density matrices

Physical Review A   May 23, 2016

Authors: Huub van Dam

Density functional theory is currently the most widely applied method in electronic structure theory. The Kohn-Sham method, based on a fictitious system of noninteracting particles, is the workhorse of the theory. The particular form of the Kohn-Sham wave function admits only idempotent one-electron density matrices whereas wave functions of correlated electrons in post-Hartree-Fock methods invariably have fractional occupation numbers. Here we show that by generalizing the orbital concept and introducing a suitable dot product as well as a probability density, a noninteracting system can be chosen that can represent the one-electron density matrix of any system, even one with fractional occupation numbers. This fictitious system ensures that the exact electron density is accessible within density functional theory. It can also serve as the basis for reduced density matrix functional theory. Moreover, to aid the analysis of the results the orbitals may be assigned energies from a mean-field Hamiltonian. This produces energy levels that are akin to Hartree-Fock orbital energies such that conventional analyses based on Koopmans' theorem are available. Finally, this system is convenient in formalisms that depend on creation and annihilation operators as they are trivially applied to single-determinant wave functions.

Fault Modeling of Extreme Scale Applications using Machine Learning

Proc. IPDPS 2016   May 24, 2016

Authors: Abhinav Vishnu, Huub van Dam, Nathan Tallent, Darren Kerbyson, Adolfy Hoisie

Faults are commonplace in large scale systems. These systems experience a variety of faults such as transient, permanent and intermittent. Multi-bit faults are typically not corrected by the hardware resulting in an error. This paper attempts to answer an important question: Given a multi-bit fault in main memory, will it result in an application error — and hence a recovery algorithm should be invoked — or can it be safely ignored? We propose an application fault modeling methodology to answer this question. Given a fault signature (a set of attributes comprising of system and application state), we use machine learning to create a model which predicts whether a multibit permanent/transient main memory fault will likely result in error. We present the design elements such as the fault injection methodology for covering important data structures, the application and system attributes which should be used for learning the model, the supervised learning algorithms (and potentially ensembles), and important metrics. We use three applications — NWChem, LULESH and SVM — as examples for demonstrating the effectiveness of the proposed fault modeling methodology.

A comparison of different methods to implement higher order derivatives of density functionals

BNL-112219-2016-IR   May 2016

Authors: Huub van Dam

Density functional theory is the dominant approach in electronic structure methods today. To calculate properties higher order derivatives of the density functionals are required. These derivatives might be implemented manually, by automatic differentiation, or by symbolic algebra programs. Different authors have cited different reasons for using the particular method of their choice. This paper presents work where all three approaches were used and the strengths and weaknesses of each approach are considered. It is found that all three methods produce code that is sufficiently performant for practical applications, despite the fact that our symbolic algebra generated code and our automatic differentiation code still have scope for significant optimization. The automatic differentiation approach is the best option for producing readable and maintainable code.

Education

Utrecht University

PhD, Theoretical Chemistry, 1993 - 1997

Radboud University Nijmegen

Msc, Chemistry, 1990 - 1993

Bsc, Chemical Engineering, 1985 - 1989

Organizations

American Chemical Society (ACS)

February 2011 to Present

Huub van Dam

Application Architect at Brookhaven National Laboratory

huubvandam@aol.com

Contact Huub on LinkedIn